




Three Distinct Proteases Are Responsible for Overall Cell Surface Proteolysis in *Streptococcus thermophilus*

Mylène Boulay,^a Coralie Metton,^{a*} Christine Mézange,^a Lydie Oliveira Correia,^a Thierry Meylheuc,^{a§} Véronique Monnet,^a Rozenn Gardan,^a  Vincent Juillard^a

^aUniversité Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, Jouy-en-Josas, France

ABSTRACT The lactic acid bacterium *Streptococcus thermophilus* was believed to display only two distinct proteases at the cell surface, namely, the cell envelope protease PrtS and the housekeeping protease HtrA. Using peptidomics, we demonstrate here the existence of an additional active cell surface protease, which shares significant homology with the SepM protease of *Streptococcus mutans*. Although all three proteases—PrtS, HtrA, and SepM—are involved in the turnover of surface proteins, they demonstrate distinct substrate specificities. In particular, SepM cleaves proteins involved in cell wall metabolism and cell elongation, and its inactivation has consequences for cell morphology. When all three proteases are inactivated, the residual cell-surface proteolysis of *S. thermophilus* is approximately 5% of that of the wild-type strain.

IMPORTANCE *Streptococcus thermophilus* is a lactic acid bacterium used widely as a starter in the dairy industry. Due to its “generally recognized as safe” status and its weak cell surface proteolytic activity, it is also considered a potential bacterial vector for heterologous protein production. Our identification of a new cell surface protease made it possible to construct a mutant strain with a 95% reduction in surface proteolysis, which could be useful in numerous biotechnological applications.

KEYWORDS *Streptococcus thermophilus*, cell surface proteolysis, peptidomics, PrtS, HtrA, SepM

Streptococcus thermophilus is the only species of genus *Streptococcus* that is generally recognized as safe (GRAS) (1). This lactic acid bacterium (LAB) is used widely in the dairy industry, as both a key partner for yogurt fermentations and a cheese starter in various cheese technologies (2). Like other LABs, its growth in protein-containing media, such as milk, depends on protein degradation, which provides it with the amino acids necessary for growth. For this reason, the proteolytic system of LABs has been studied intensively for many years (3). The first step of the casein utilization process is extracellular casein hydrolysis, which in *S. thermophilus* is mediated by the cell envelope protease PrtS (located in the chromosome) (4) and in *Lactococcus lactis* by the homolog PrtP (located on a plasmid) (5). These serine proteases are anchored to the cell wall via a sortase-dependent mechanism. For many years, they were thought to be the sole surface proteases of LABs, as their absence prevented casein degradation and thus affected growth in milk. To date, no other function than nitrogen supply has been attributed to PrtS/PrtP in these two species, although a PrtS homolog, the challisin protease of *Streptococcus gordonii*, has been shown to degrade the *S. gordonii* competence stimulating peptide (CSP), which is responsible for the induction of competence for natural transformation (6).

More recently, a second cell surface serine protease, HtrA, was characterized in *L. lactis* (7). HtrA is a membrane-bound protease that faces the external medium; it is not thought to be involved in nutrition since it has not been shown to participate in

Citation Boulay M, Metton C, Mézange C, Oliveira Correia L, Meylheuc T, Monnet V, Gardan R, Juillard V. 2021. Three distinct proteases are responsible for overall cell surface proteolysis in *Streptococcus thermophilus*. *Appl Environ Microbiol* 87:e01292-21. <https://doi.org/10.1128/AEM.01292-21>.

Editor Johanna Björkroth, University of Helsinki

Copyright © 2021 American Society for Microbiology. All Rights Reserved.

Address correspondence to Vincent Juillard, vincent.juillard@inrae.fr.

*Present address: Coralie Metton, DuPont Nutrition and Health, BP10, Dangé-St-Romain, France.

§Present address: Thierry Meylheuc, UMR SQPOV, INRAE, Site Agroparc CS40509, Avignon, France.

Received 29 June 2021

Accepted 13 September 2021

Accepted manuscript posted online 22 September 2021

Published 10 November 2021

the casein degradation process. HtrA is conserved widely in bacteria and eukaryotic organisms (8). As is the case in other microorganisms, its main function in *L. lactis* is the degradation of misfolded proteins that accumulate under conditions of thermal stress. Lactococcal HtrA is also involved in the processing of certain exported proteins, such as autolysins (7). More widely, HtrA plays a role in many regulatory mechanisms. For example, in *Escherichia coli*, HtrA cleaves the periplasmic domain of RseA; this step is followed by a cascade of proteolytic events that ultimately results in the induction of specific stress response genes (8). Other examples have been identified in *Streptococcus pneumoniae*, in which HtrA represses competence for natural transformation through the selective degradation of DNA uptake machinery (9), and the CSP pheromone, which initiates competence and bacteriocin production (10, 11). To date, no similar regulatory roles have been attributed to HtrA in LABs. Of the several other proteases that have been characterized in *L. lactis* and *S. thermophilus*, all are located inside the cell; no other cell surface proteases with activity outside the cell (i.e., shed-dases) have been identified so far in these bacteria.

Recently, there has been interest in broadening the applications of LABs beyond their traditional uses in the food industry (12). Due to their GRAS status, *L. lactis* and more recently *S. thermophilus* have been proposed as potential bacterial vectors for heterologous protein production (13–15). Among the advantages presented by these two species is the fact that they are regarded as poorly proteolytic, due to the low number of cell surface proteases they produce compared with *Bacillus subtilis*, for instance (16). In previous work, we developed a peptidomic approach for analyzing the extracellular proteolytic activity of a PrtP-deficient strain of *L. lactis* (17). Surprisingly, we found that HtrA was responsible for the release of only a portion of the peptides accumulating in the supernatant, suggesting that uncharacterized protease(s) other than HtrA were active at the cell surface. The aim of the present study was thus to identify these unknown proteases in order to ultimately obtain a strain deprived fully of surface proteolytic activity. We decided to use *S. thermophilus* rather than *L. lactis* as the study bacterium for a technical reason; although the two species are closely related (1), the conditions for entering a state of competence for natural transformation are known for *S. thermophilus* but not for *L. lactis* (18–21). It is thus much easier to construct mutant strains of *S. thermophilus* than those of *L. lactis*. We first applied our powerful peptidomic approach to confirm the presence of an as-yet-unknown shed-dase in *S. thermophilus*, as our previous study had demonstrated in *L. lactis*. This effort revealed an additional protease at the cell surface of *S. thermophilus* that is orthologous to the SepM protease of *Streptococcus mutans* (22). The three sheddases PrtS, HtrA, and SepM displayed different substrate specificities. None of them seemed to be involved in pheromone maturation, but all three played a role in the turnover of surface proteins, including PrtS itself. Moreover, SepM seemed to be involved in cell division processes. Through the construction of single- and multiple-mutant strains, we were able to determine that together PrtS, HtrA, and SepM account for 95% of the cell surface proteolytic activity of *S. thermophilus*.

RESULTS

The extracellular peptidome of *S. thermophilus* LMD9 results mainly from the hydrolysis of surface proteins. Peptides accumulating in the growth medium of *S. thermophilus*—collectively referred to as the exopeptidome—were identified using a two-dimensional liquid chromatography-mass spectrometry (LC-MS) approach. The exopeptidome of *S. thermophilus* LMD9 contained 650 distinct oligopeptides, with each composed of between 7 and 32 amino acids (mean of 3 independent biological repetitions; SD, 161) (see Table S1 in the supplemental material). One-third of them (188) were found systematically in each experiment, and more than one-half of the peptides were identified in at least 2 experiments. This degree of reproducibility is similar to what has already been described for the exopeptidome of *L. lactis* (17). In *S. thermophilus*, the exopeptidome resulted from the hydrolysis of 215 distinct proteins (SD, 25) (Table S1). Most of the peptides (66%) present were generated via the hydrolysis of

TABLE 1 Bacterial strains used in the present study

Strain	Genotype	Reference
LMD9	Wild type	57
LMD9 Δ htrA	STER_RS09790::aphA3	This study
LMD9 Δ prtS	STER_RS04165::erm	This study
LMD9 Δ sepM	STER_RS07910::spec	This study
LMD9 Δ htrA Δ prtS	STER_RS09790::aphA3, STER_RS04165::erm	This study
LMD9 Δ htrA Δ sepM	STER_RS09790::aphA3, STER_RS07910::spec	This study
LMD9 Δ prtS Δ sepM	STER_RS04165::erm, STER_RS07910::spec	This study
LMD9 Δ htrA Δ prtS Δ sepM	STER_RS09790::aphA3, STER_RS04165::erm, STER_RS07910::spec	This study
CNRZ1066	Wild type	63
CNRZ1066 Δ htrA Δ sepM	STR_RS09505::aphA3, STR_RS07745::spec	This study

surface-located or secreted proteins, according to LocateP database predictions (23). Other peptides originated from transmembrane or cytoplasmic proteins, with some probably present in multiple locations (24).

Previous work in *S. thermophilus* has characterized two distinct proteases that could be responsible for the formation of the exopeptidome, namely, PrtS (STER_RS04165) and HtrA (STER_RS09790). To evaluate this role, we constructed two single-deletion mutants (Table 1) and analyzed their respective exopeptidomes. Inactivation of the *prtS* gene resulted in a slight decrease (21%) in the number of peptides identified, while inactivation of the *htrA* gene had a larger effect (43% reduction) (Table S1). In both cases, the reduction in the total number of peptides was the effect of a decrease in the amount of peptides released from surface proteins, consistent with the cellular localization of the proteases (Fig. 1). Furthermore, the number of peptides released from the PrtS protein was significantly reduced in the exopeptidome of the Δ htrA mutant strain (3 peptides versus 32 in the wild-type strain, $P = 0.0463$), indicating that HtrA actively cleaves PrtS during the growth of *S. thermophilus*. In contrast, the number of peptides released from transmembrane or cytoplasmic proteins was not significantly modified in the absence of PrtS or HtrA (Fig. 1).

The inactivation of both *htrA* and *prtS* led to a slightly larger reduction (50%) in the number of peptides present in the exopeptidome (Fig. 1), with 331 distinct peptides

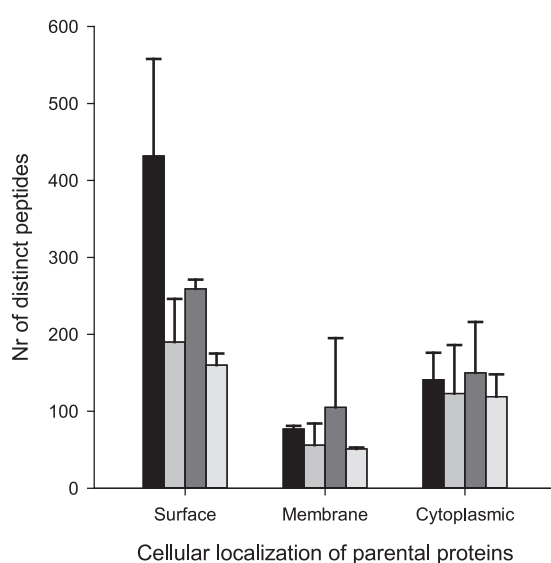


FIG 1 The number of distinct peptides identified in the supernatant of strains of *S. thermophilus* LMD9, as a function of the cellular compartment of the parental protein. The identity of each group of bars, from left to right, is as follows: LMD9, LMD9 Δ htrA, LMD9 Δ prtS, and LMD9 Δ htrA Δ prtS. Values are the mean of three biological replicates, with standard deviation of the mean. *, $P < 0.05$.

detected (mean of 3 independent experiments; SD, 40). As observed with the single-mutation strains, this reduction was due mainly to a decrease in the number of peptides released from surface proteins. For instance, the two transporter substrate-binding proteins STER_RS04220 and STER_RS07145 were significantly less degraded in the double mutant than in the wild-type strain (P values of 0.0369 and 0.0463, respectively) (Table S1). However, this decrease in peptide count was not statistically significant in the absence of one protease only (in the absence of HtrA and PrtS, P values were 0.1266 and 0.1266 for STER_RS04220 and 0.09896 and 0.5066 for STER_RS07145, respectively). This finding suggests the existence of a synergism between HtrA and PrtS in the turnover of these surface proteins.

The gene *STER_RS07910* encodes a functional cell surface protease. From these results, it was very clear that proteolytic activity was still present at the cell surface of the double mutant strain LMD9 Δ htrA Δ prtS. We thus developed an *in silico* approach to identify candidates for this new surface protease. According to the MEROPS database, release 10.0 (<http://merops.sanger.ac.uk/cgi-bin/speccards?sp=sp000991;type=peptidase;strain=498>), the genome of *S. thermophilus* LMD9 encodes 45 members of protease-peptidase families, which represents 2.6% of the total number of genes of this strain (25). An analysis of the annotation of the *S. thermophilus* genome annotation (<https://www.ncbi.nlm.nih.gov/genome/genomes/4207>) revealed the possible presence of additional proteases. In total, 52 proteins were predicted to contain a proteolytic domain (Table 2), and of these, 11 were predicted to be located at the cell surface. Two (PrtS and HtrA) have already been implicated in the formation of the exopeptidome, and it seemed very likely that at least one of the other nine proteins also participates in peptide accumulation. Interestingly, peptides released from five of these nine putative surface proteases (STER_RS00770, RS01270, RS05675, RS07910, and RS08505) were detected in the *S. thermophilus* exopeptidomes; this finding indicated that these putative proteases were produced efficiently during growth (Table S1) and therefore represent potential candidates. In contrast, the four other proteases were apparently not expressed under these experimental conditions and therefore did not appear to be relevant candidates. Of the five candidate proteases, STER_RS00770 and RS01270 are annotated as D-Ala-D-Ala carboxypeptidases and are thought to be involved mainly in peptidoglycan formation (26); they thus probably did not represent the best candidates for contributing to the exopeptidome. Likewise, STER_RS08505 is a signal peptide peptidase and would be unlikely to make a significant contribution to the exopeptidome. The two remaining candidates were STER_RS05675 and STER_RS07910.

Our earlier analysis of *L. lactis* IL1403 had suggested the existence of a third active cell surface protease, other than PrtP and HtrA, that participated in the formation of the exopeptidome in that species (17). We thus decided to start from the most parsimonious hypothesis, namely, that this third cell surface protease is homologous between the two closely related bacteria. Of our two most likely candidates, STER_RS05675 had no counterpart in the genome of *L. lactis* IL1403, while STER_RS07910 did. We therefore created a deletion mutant for STER_RS07910 and analyzed its exopeptidome (Table S1). The results indicated that STER_RS07910 contributed actively to exopeptidome formation; the number of peptides detected in the supernatant of the mutant strain ($n = 324$; mean of 3 assessments; SD, 100) was 2 times lower than that of the wild-type strain. Once again, this decrease was the consequence of a significant reduction in surface protein hydrolysis, especially of lipoproteins and membrane-anchored proteins ($P = 0.04953$, for both classes). In contrast to HtrA and PrtS, STER_RS07910 did not appear to cleave proteins located in the cell wall in any significant way, as the number of these peptides released was in the same range as that found in the exopeptidome of the wild-type strain (Fig. 2).

To determine if PrtS, HtrA, and STER_RS07910 are responsible for the majority of exopeptidome formation—or if, instead, additional surface protease(s) remained to be identified—a mutant strain deficient in these three sheddases was constructed, and its exopeptidome was characterized (Table S1). In the absence of these 3 proteases, the

TABLE 2 Predicted proteases of *S. thermophilus* LMD9

Family ^a	Peptidase or homologue (subtype) ^b	Cellular localization (locate P)	<i>S. thermophilus</i> LMD9	
			Gene	Length (aa) ^c
S11	Serine-type D-Ala-D-Ala carboxypeptidase	Secreted	STER_RS00565	415
nd	D-Ala-D-Ala carboxypeptidase	Secreted (secretome P)	STER_RS03750	331
S8	Subtilisin-like serine protease (PrtS)	Cell wall anchored	STER_RS04165	1,618
M15	D-Ala-D-Ala carboxypeptidase	Membrane anchored	STER_RS00770	265
nd	D-Ala-D-Ala carboxypeptidase	Membrane anchored	STER_RS01270	775
S26	Signal peptidase I (sipB)	Membrane anchored	STER_RS05430	185
M10	Zn-dependent protease	Membrane anchored	STER_RS05675	239
C60	Sortase A	Membrane anchored	STER_RS06195	762
S16	Predicted protein-containing PDZ and Lon-protease domains	Membrane anchored	STER_RS07910	345
S36	Signal peptidase I (sipA)	Membrane anchored	STER_RS08505	207
S1	HtrA peptidase	Membrane anchored	STER_RS09790	411
M41	Cell division protein FtsH	Multitransmembrane	STER_RS00070	655
M50	RIP metalloprotease RseP (eep)	Multitransmembrane	STER_RS01205	420
A8	Signal peptidase II	Multitransmembrane	STER_RS02730	153
M48	Zinc metalloprotease HtpX	Multitransmembrane	STER_RS03715	299
A24	Type-4 prepilin-like protein specific leader peptidase	Multitransmembrane	STER_RS03755	215
S33	SCO7095-type peptidase	Multitransmembrane	STER_RS06090	851
C39	Bacteriocin-processing peptidase	Multitransmembrane	STER_RS08100	717
S54	Rhomboid family intramembrane serine protease	Multitransmembrane	STER_RS08850	224
M79	CAAX amino protease	Multitransmembrane	STER_RS08935	220
C44	Family C44 unassigned peptidases	Cytoplasmic	STER_RS00260	479
M29	PepS aminopeptidase	Cytoplasmic	STER_RS00485	413
C1	Aminopeptidase C	Cytoplasmic	STER_RS01340	445
S14	ATP-dependent Clp protease proteolytic subunit	Cytoplasmic	STER_RS01925	196
M3	Oligoendopeptidase F	Cytoplasmic	STER_RS02405	601
T5	Arginine biosynthesis bifunctional protein	Cytoplasmic	STER_RS02460	397
M20	Peptidase M20	Cytoplasmic	STER_RS02920	381
nd	Endoribonuclease ^d	Cytoplasmic	STER_RS03285	165
M24	Proline dipeptidase	Cytoplasmic	STER_RS03340	361
U32	Peptidase U32	Cytoplasmic	STER_RS03560	309
U32	Protease	Cytoplasmic	STER_RS03565	428
C44	Family C44 unassigned peptidases	Cytoplasmic	STER_RS04470	602
C56	Putative intracellular protease/amidase	Cytoplasmic	STER_RS04770	182
M1	Aminopeptidase N	Cytoplasmic	STER_RS04990	846
M20	Dipeptidase	Cytoplasmic	STER_RS05375	468
M20	Tripeptidase PepT	Cytoplasmic	STER_RS05425	407
C26	Glutamine amidotransferase	Cytoplasmic	STER_RS05930	202
M3	Oligoendopeptidase F	Cytoplasmic	STER_RS06400	777
C26	Glutamine amidotransferase (class I)	Cytoplasmic	STER_RS07030	231
M24	Methionine aminopeptidase	Cytoplasmic	STER_RS07460	286
C26	Glutamine amidotransferase	Cytoplasmic	STER_RS07635	188
S12	Esterase EstB	Cytoplasmic	STER_RS07995	318
nd	C3-glycoprotein-degrading proteinase ^e	Cytoplasmic	STER_RS08000	251
S15	Xaa-Pro dipeptidyl peptidase	Cytoplasmic	STER_RS08015	755
M24	Xaa-Pro aminopeptidase	Cytoplasmic	STER_RS08390	353
M22	Glycoprotein endopeptidase	Cytoplasmic	STER_RS08535	228
S16	DNA repair protein RadA	Cytoplasmic	STER_RS08805	415
M20	Metal-dependent carboxypeptidase	Cytoplasmic	STER_RS08860	377
M42	Glutamyl aminopeptidase	Cytoplasmic	STER_RS08925	355
M13	Oligoendopeptidase O3	Cytoplasmic	STER_RS09100	631
M16	Peptidase M16	Cytoplasmic	STER_RS09720	425
M16	Peptidase M16	Cytoplasmic	STER_RS09725	416

^and, not determined.^bClassification based on the MEROPS database (release 10.0, <http://merops.sanger.ac.uk>), enriched with *S. thermophilus* strain annotation.^caa, amino acids.^dAnnotated as a metalloprotease in some *S. thermophilus* strains (MTH17CL396; MN-ZLW-002; TH435,436,982,985,1477; ASCC1275; 1F8TC).^eContains domains CppA_N and _C, N- and C-terminal domains of the CppA protein, respectively. CppA is found in species of *Streptococcus*. CppA is a putative C3-glycoprotein-degrading proteinase, involved in pathogenicity.

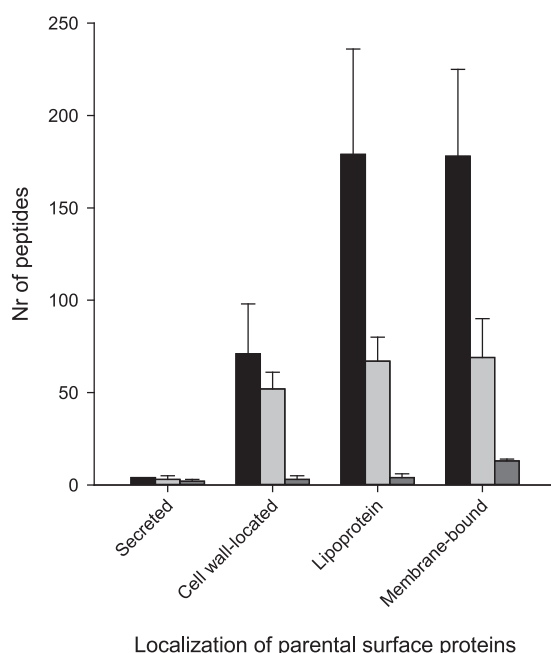


FIG 2 The number of peptides released from surface proteins. Black, medium, and dark gray bars represent LMD9, LMD9ΔSTER_RS07910, and LMD9ΔhtrAΔprtSΔSTER_RS07910 strains, respectively. Values are the mean of three biological replicates, with standard deviation of the mean. *, $P < 0.05$.

number of peptides released from surface proteins was reduced 25-fold, with only 21 distinct peptides detected (Fig. 2). A comparison of the number of surface peptides in the exopeptidome of the double mutant ΔhtrAΔprtS and the triple mutant ΔhtrAΔprtSΔSTER_RS07910 revealed that inactivation of STER_RS07910 induced a 9-fold reduction in the number of peptides released from surface proteins. In contrast, the number of peptides released from transmembrane ($n = 34$; SD, 4) and cytoplasmic proteins ($n = 71$; SD, 21) was affected only slightly in the triple mutant compared with the wild-type strain. These results indicated clearly that these three proteins account for the majority of cell surface proteolysis. Based on the number of surface peptides detected in the medium (spectral counting), we estimated that the residual surface proteolysis of the triple mutant strain represented only 5% of that of the wild-type strain. We further verified that the modification of the extracellular peptidome was not due to poor growth of the strain and found that the absence of the three proteases did not affect growth markedly in chemically defined medium (CDM) (see Fig. S1 in the supplemental material).

STER_RS07910 is conserved widely. The presence of PrtS at the cell surface of *S. thermophilus* is strain dependent (4, 27); in the 151 distinct genomic sequences of *S. thermophilus* that are available, only 30% have the *prtS* gene. Instead, HtrA and STER_RS07910 proteins are present in 100% and 99% of the strains, respectively, with sequence identities systematically higher than 90% with the LMD9 sequences (Fig. 3). Strains of *S. thermophilus* can therefore be clustered into two classes based on their assemblage of surface proteases; a minority of strains (less than one-third) contain three surface proteases, while the majority of the strains (more than two-thirds) have only two surface proteases, namely, HtrA and STER_RS07910. To ascertain that these two proteases are still responsible for the essential activity of surface proteolysis in the latter group, we repeated our assessments using strain CNRZ1066, which naturally lacks PrtS. Based on the spectral counting of peptides released from surface proteins, inactivation of the 2 protease-encoding genes *htrA* and *ster_RS07910* resulted in a 94% loss of surface proteolysis, with spectral counts of 240 peptides for the wild-type and 15 peptides for the double mutant strain.

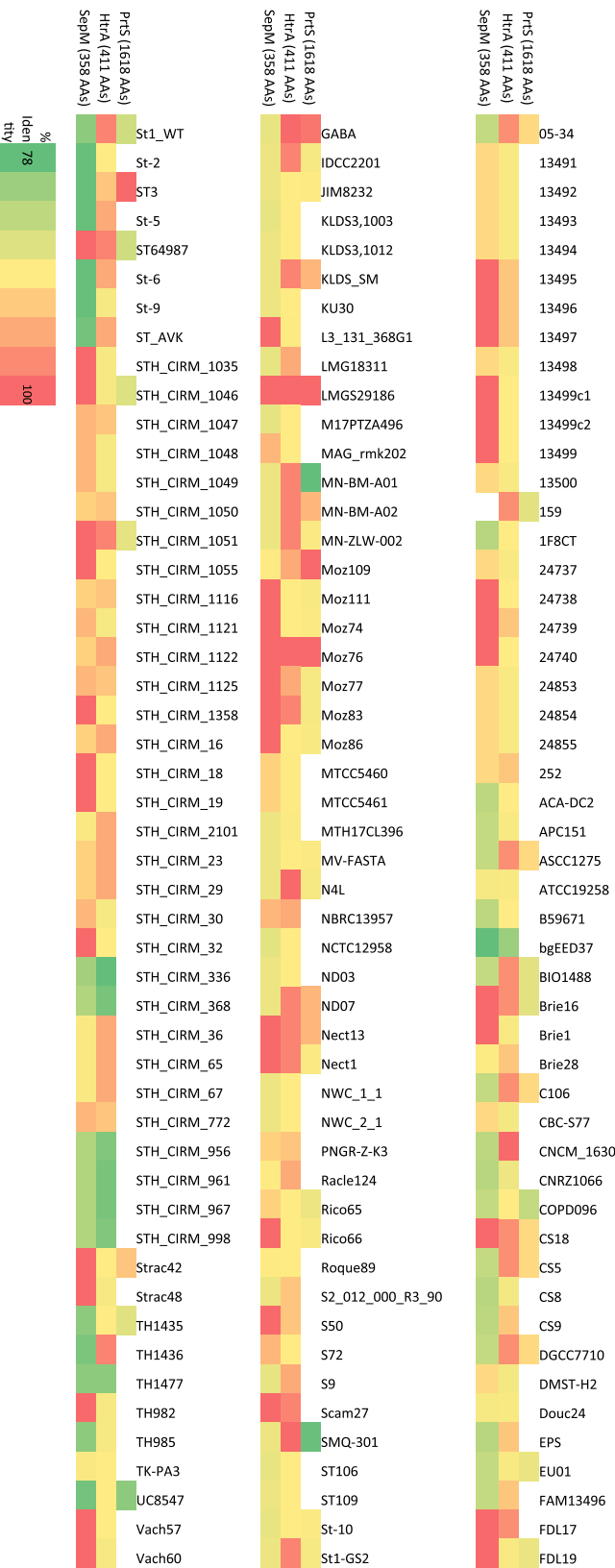


FIG 3 Conservation of the three cell surface proteases in *S. thermophilus*. LMD9 is used as the reference strain.

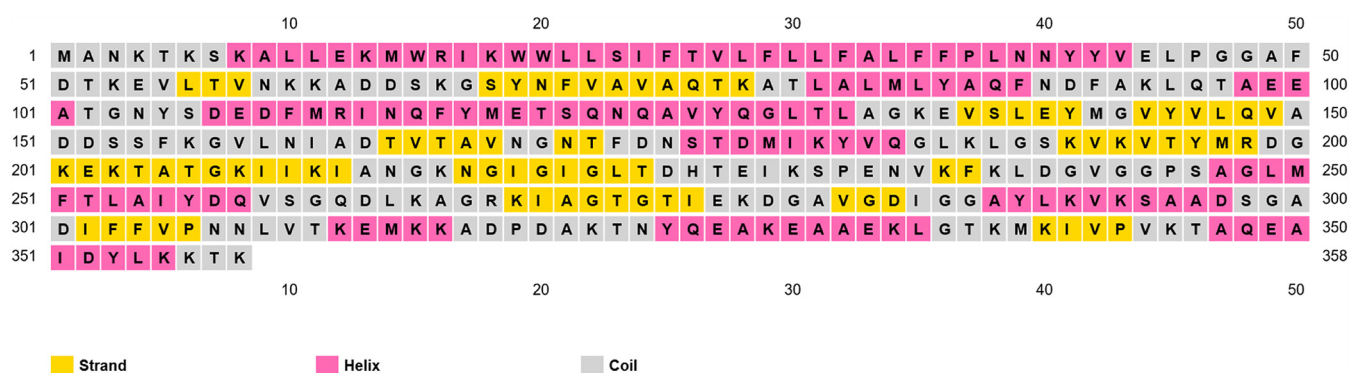


FIG 4 SepM secondary structure prediction. Prediction was obtained from the PSIPRED Web server version 4.0 (<http://bioinf.cs.ucl.ac.uk/psipred>).

STER_RS07910 belongs to the peptidase S16 subfamily and is a member of the Clusters of Orthologous Groups of proteins (COG) family of signal transduction mechanisms. It is predicted to contain a transmembrane domain at the N terminus of the protein (position 19 to 36), a PDZ-domain of ATP-dependent Lon serine protease (position 130 to 193) and a C-terminal proteolytic domain of Lon protease (position 231 to 341). In contrast to most Lon proteases, STER_RS07910 does not have the classical walker A and B motifs responsible for ATP fixation. The family type peptidase is the Lon-A peptidase from *Escherichia coli* (MEROPS accession MAR0000485 [<https://www.ebi.ac.uk/merops/cgi-bin/famsum?family=S16>]), with the active site residues S₆₇₉ and K₇₂₂ (28). These two residues are conserved in STER_RS07910 (S₂₄₆ and K₂₉₁) and most probably constitute the active site also in this protease. STER_RS07910 is conserved in streptococci and, in particular, is highly conserved with SepM of *Streptococcus mutans* (60.6% identity, 74.9% similarity). For that reason, STER_RS07910 will be renamed SepM throughout the remaining text. The 2D structure of SepM was predicted using the PSIPRED protein structure prediction tool (<http://bioinf.cs.ucl.ac.uk/psipred>) (29). SepM is predicted to have several helix and strand regions (Fig. 4) and, therefore, according to the CATH classification (<https://www.cathdb.info>), belongs to the $\alpha + \beta$ class of proteins (30).

We analyzed the organization of *sepM* flanking genes in the 84 circularized genomes of *S. thermophilus* and noted a high degree of conservation. Most of the strains (78 of 84) have the same organization as LMD9, whereas 6 strains (CIRM336, 368, 956, 961, 967, and 998) have pseudogenes (IS3 family transposase derived) downstream of *sepM* (Fig. 5). In addition, we studied the location of *sepM* on the chromosome of these strains. For 77 of the 84 strains, the *sepM* gene is located at 47 minutes (SD, 6.7) in the chromosome. However, in 7 strains (CIRM23, 32, 956, 961, EPS, MAGrmk202, and TK-PA3), *sepM* is located in the first 10 minutes of the chromosome. Predictions from the BioCyc database collection suggested that *sepM* is cotranscribed with *coaD* and *rsmD* (31), and this prediction was supported by transcriptional data obtained with *S. thermophilus* LMD9 (32) or N4L (33). These data indicate that *sepM* is the last gene on an operon made up of

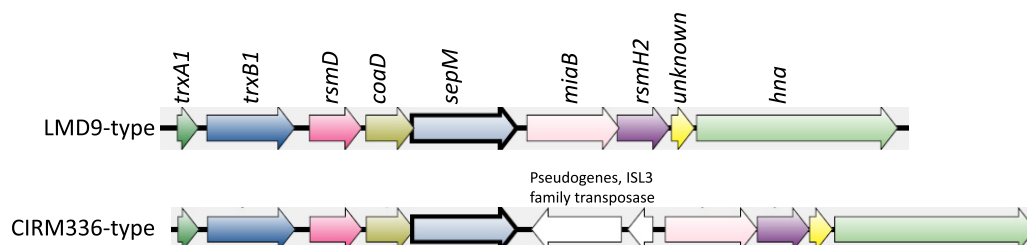


FIG 5 *sepM*-flanking genes in *S. thermophilus*. Synteny prediction was obtained from the SyntTax Web server (<https://archaea.i2bc.paris-saclay.fr/SyntTax>).

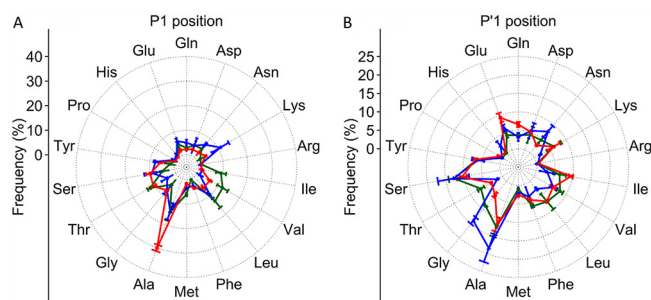


FIG 6 Amino acid frequencies at the cleavage sites of the peptides. (A) P1 position of the cleavage site. (B) P'1 position of the cleavage site. Blue line, strain LMD9Δ*htrA*Δ*sepM*, producing only PrtS; green line, strain LMD9Δ*prtS*Δ*sepM*, producing only HtrA; red line, strain LMD9Δ*htrA*Δ*prtS*, producing only SepM.

three genes with no apparent functional relationship among them; *rsmD* encodes a methyltransferase involved in ribosome biogenesis and *coaD* encodes a transferase belonging to the pantothenate and coenzyme A (CoA) biosynthesis pathway. According to work by Eng and colleagues, *sepM* is not considered an atypical gene acquired by horizontal gene transfer (34).

At the protein level, sequences encoded by the *sepM* flanking genes are extremely well conserved within *S. thermophilus*. In all streptococci that have been analyzed, the proteins encoded in the upstream region of *sepM* are conserved, while the downstream region is conserved only in some (see Table S2 in the supplemental material).

The three sheddases have distinct substrate specificities. We evaluated the substrate specificity of the three surface proteases by constructing double mutant strains. Our assumption was that the mutant strains LMD9Δ*htrA*Δ*prtS*, LMD9Δ*htrA*Δ*sepM*, and LMD9Δ*prtS*Δ*sepM* would each contain only one cell surface protease, namely, SepM, PrtS, or HtrA, respectively. We then assessed the specificity of the cleavage of each surface protease based on the exopeptidome composition (unique peptides) of the corresponding double mutant. We first identified the amino acids flanking the N- and C-terminal cleaved peptide and then calculated their relative frequencies. These results indicated that the three sheddases do not have the same specificities of cleavage (Fig. 6). In the P1 position (corresponding to the N-side of the cleaved bond), PrtS has a more pronounced preference for hydrophilic charged amino acids than SepM, HtrA has a stronger preference for hydrophobic branched-chain amino acids than both SepM and PrtS, and SepM has a much more marked preference for Ala than PrtS and HtrA. Differences in the specificities of the three proteases were also noted in the P'1 position, which corresponds to the C-terminal side of the cleaved bond.

An analysis of the degraded proteins (based on peptides identified in the exopeptidome) also revealed differences among the sheddases regarding the proteins they cleave (see Table S3 in the supplemental material). In the wild-type strain, the most intensively cleaved surface proteins were two cell wall-associated proteins (STER_RS00970 and STER_RS04165), six lipoproteins (STER_RS00685, STER_RS04220, STER_RS04960, STER_RS06930, STER_RS06940, and STER_RS07945), and four membrane-bound proteins (STER_RS01345, STER_RS05805, STER_RS07910, and STER_RS08990). The bifunctional 2',3'-cyclic nucleotide 2'-phosphodiesterase/3'-nucleotidase precursor protein (STER_RS00970) is actively cleaved by HtrA and poorly cleaved by SepM and PrtS. Its spectral count strongly decreased in the exopeptidomes of LMD9Δ*htrA*, LMD9Δ*htrA*Δ*prtS*, and LMD9Δ*htrA*Δ*sepM*; remained unchanged in LMD9Δ*sepM*; and was reduced only slightly in LMD9Δ*prtS* and LMD9Δ*prtS*Δ*sepM*. The cell wall-associated protein PrtS (STER_RS04165) was cleaved very efficiently by HtrA but cleaved only weakly by SepM. A similar trend was observed with another (less intensively cleaved) cell wall-associated protein (STER_RS00210, CHAP domain-containing protein). Together, these results suggest that SepM is only weakly active on cell wall-associated proteins. Conversely, HtrA

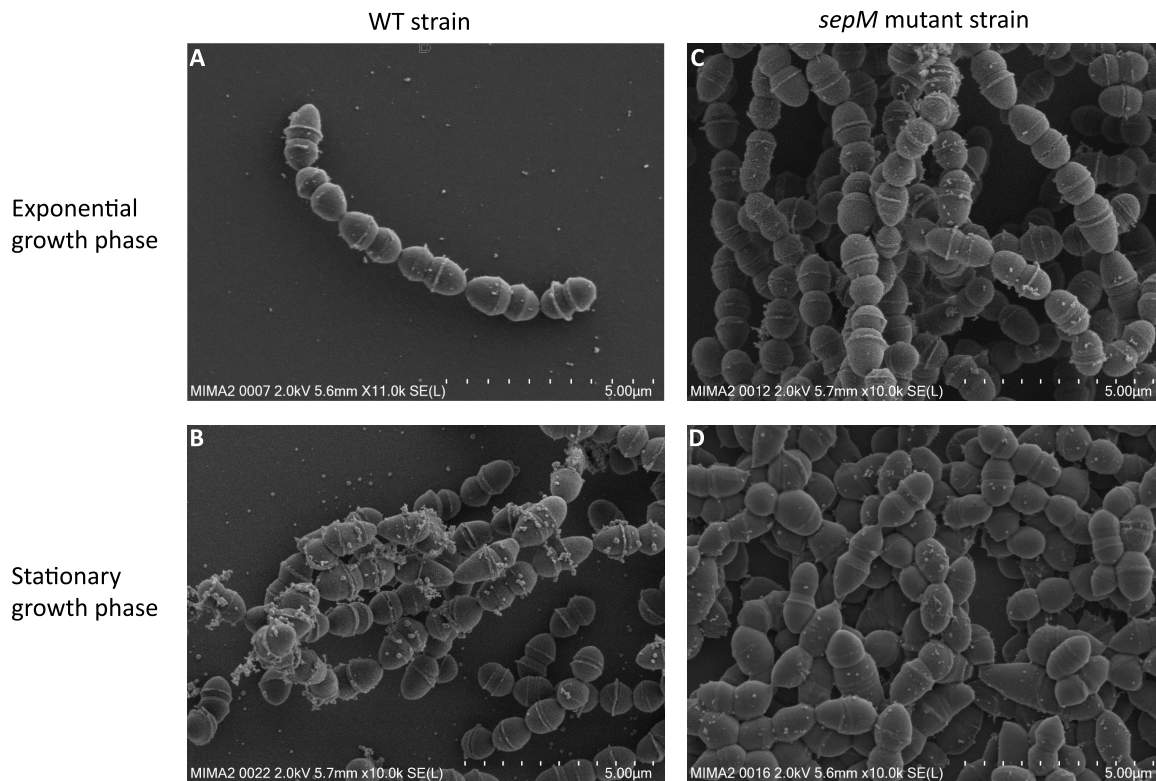


FIG 7 Scanning electron microscopy of *S. thermophilus* LMD9. Wild type (A, C) and $\Delta sepM$ (B, D) strains in exponential (A, B) and stationary (C, D) phases of growth.

had no significant action on the amino acid ABC transporter periplasmic protein STER_RS07945, whereas SepM was highly efficient. The late competence protein STER_RS08990 involved in DNA uptake was cleaved by both HtrA and SepM but not by PrtS. All three proteases were able to degrade the penicillin-binding protein STER_RS01345.

When we specifically examined the proteins cleaved by SepM, we found that a significant proportion were involved in cellular division or cell wall metabolism (e.g., PBP-1A, PBP-2A, PBP-2B, MesJ, MreC, PG hydrolases, and LCP proteins). We therefore decided to compare the morphology of the wild type and mutant using scanning electron microscopy (Fig. 7). Cells were compared both at the exponential and stationary phases of growth, and similar trends were observed. Wild-type cells had a regular shape, and some particles, presumably extracellular vesicles, were observed at the cell surface. In the mutant cells, such particles were present at a higher frequency. Moreover, the shape of the mutant was sometimes affected, with unusual division planes. Abnormal cell shape was more pronounced at the stationary phase than in the exponential growth phase.

Role of surface proteases in quorum sensing. As mentioned above, SepM belongs to the COG family of signal transduction mechanisms. In *S. mutans*, SepM is responsible for the maturation of the competence stimulating peptide (CSP) and is therefore involved in quorum sensing (22). The three surface proteases differ in their involvement in this activity: HtrA degrades the CSP of *S. pneumoniae* but not that of *S. mutans* (10, 11, 35), whereas the *S. gordonii* homolog of PrtS cleaves the CSP of *S. mutans* (6). Taking this context into account, we assessed a possible role for these three sheddases in quorum sensing in *S. thermophilus*. This bacterium is known to express distinct quorum sensing systems, including streptide and bacteriocin production (36–38) and natural competence (18, 19). We first studied the production of streptide, a cyclic nonapeptide (AK*GDGW*KVM) that accumulates during the growth of *S. thermophilus* as the result of a quorum sensing mechanism;

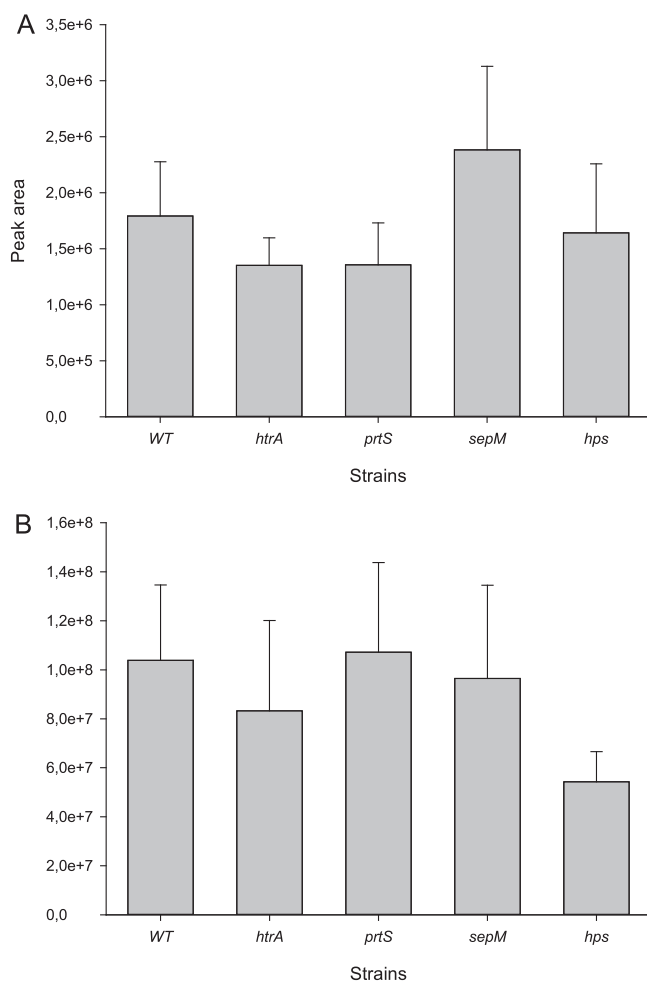


FIG 8 Quantification of the SHP₁₃₅₈ pheromone (A) and streptide (B) in the supernatant of LMD9 wild-type and mutant strains by LC-MS/MS. WT, LMD9 wild type; *htrA*, LMD9Δ*htrA*; *prtS*, LMD9Δ*prtS*; *sepM*, LMD9Δ*sepM*; *hps*, LMD9Δ*htrA*Δ*prtS*Δ*sepM*. Values are the mean of six independent determinations with SD.

the signaling peptide involved is a short hydrophobic nonapeptide, SHP₁₃₅₈ (EGIIIVVG), secreted in the external medium (37). Streptide and SHP₁₃₅₈ were detected by mass spectrometry in the supernatant of LMD9 and mutant strains grown in CDM and were quantified by integrating the area of the peaks on extracted ion chromatograms (XICs). The inactivation of surface proteases did not result in a significant reduction in SHP₁₃₅₈, as indicated by paired Wilcoxon statistical analyses (Fig. 8A). Even in the case of the triple mutant, the area of the SHP₁₃₅₈ peak remained similar to that of the wild type, indicating that none of these surface proteases played a significant role in the maturation or degradation of pheromones in the SHP family. Similarly, streptide was not affected by the inactivation of *sepM* ($P = 0.2785$) or of any other single protease gene (Fig. 8B). However, inactivation of all three proteases did seem to slightly affect streptide production, with a 40% reduction in peak area, although this change was not statistically significant ($P = 0.0625$). To clarify this last result, we evaluated streptide concentrations in the external medium using high-performance liquid chromatography (HPLC). A direct analysis of streptide by HPLC is highly challenging, as it is produced in small amounts (39) and coelutes with components of the CDM (data not shown). We therefore used a CDM deprived of aromatic amino acids and riboflavin for this analysis and performed a streptide enrichment step by solid-phase extraction prior to HPLC analysis and quantification. Similar levels of streptide were produced by both the wild type and the triple mutant strain; there

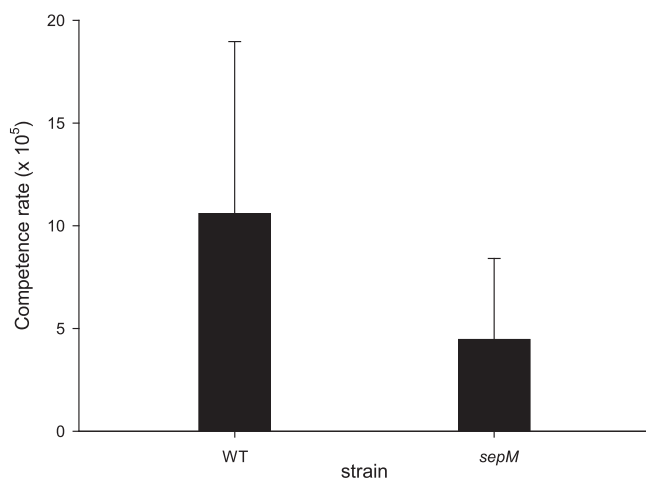


FIG 9 Transformation rate of *S. thermophilus* LMD9 wild-type and *sepM* mutant strains. Values are the mean of three independent determinations with SD. WT, LMD9 wild type; *sepM*, LMD9 Δ *sepM*. Transformation rate was determined as the ratio between the number of transformants (erythromycin-resistant colonies) and total cell count per 16.5 ng of pGhost plasmid (54).

was no difference in the peak areas obtained from four different experiments ($P = 0.4227$).

We have also evaluated the role of surface proteases with respect to natural competence for transformation in *S. thermophilus*. The development of natural competence depends on the induction of the appropriate machinery via a quorum sensing mechanism controlled by the ComS peptide pheromone (18, 19). Even though the ComS peptide is exported, its mature form does not accumulate in the supernatant of wild-type *S. thermophilus* LMD9 (40), thereby preventing its quantification by LC-MS or HPLC. For this reason, we evaluated the possible effect of *S. thermophilus* surface proteases on ComS maturation by assessing the efficiency of genetic transformation. The inactivation of *sepM* induced a slight reduction in transformation efficiency, but this was not statistically significant ($P = 0.25$) (Fig. 9). No differences were detected between mutant and wild-type strains when synthetic mature ComS was added to the reaction mixture, which suggests that SepM does not play a crucial role in ComS maturation or degradation.

DISCUSSION

One of the main uses of a peptidomic strategy is the discovery of bioactive peptides in biological fluids or food matrices (41). In the present work, we used this approach to investigate cell surface proteolysis in *S. thermophilus*. Until now, this bacterium was believed to produce only two proteases located at the cell surface, namely, PrtS and HtrA. However, when both of these proteases were inactivated, residual cell surface proteolysis was still detected, strongly suggesting the presence of an additional protease active at the cell surface. The current study identifies this previously uncharacterized sheddase as a protein presenting a high level of homology with SepM of *S. mutans*. A comparison of the exopeptidomes of wild-type, single-mutant, and multiple-mutant strains indicates that all three proteases—PrtS, HtrA, and SepM—participate in the recycling of cell surface proteins. Furthermore, no significant residual surface proteolysis was observed when these three proteases were inactivated, suggesting that they are responsible for almost all surface proteolytic activity.

HtrA and SepM are present in all strains of *S. thermophilus*, while PrtS is found in only some strains. Thus, *S. thermophilus* strains can be assigned to one of two groups based on whether they possess two or three sheddases. It seems likely that the related species *L. lactis* is similar in this regard; a previous study performed on PrtP-deficient strains of this bacterium suggested the presence of an additional cell surface protease

acting together with HtrA (17). A plausible candidate could be the membrane-bound protease YwdF, which presents significant homology with SepM (50%) and is presumably the SepM ortholog. An analysis of the exopeptidome of *L. lactis* IL1403 revealed peptides originating from YwdF (17). It therefore seems very likely that *L. lactis* can also harbor two or three distinct sheddases (HtrA, YwdF, and PrtP) responsible for overall surface proteolysis, depending on the presence of the lactose-protease plasmid in the strain (42).

HtrA, PrtS, and SepM have broad but distinct specificities, as is usually the case for serine proteases (43). In particular, based on an analysis of competence stimulating peptide (CSP) cleavage in *S. mutans*, SepM was reported to have a strict specificity, cleaving CSP between Ala18 and Leu19 residues only (44). Our analysis of a large number of cleavage sites confirmed the strong preference of SepM for Ala in the P1 position of the cleaved bond but not for Leu in the P'1 position.

All three proteases are involved in exopeptidome formation and therefore participate in the turnover of cell surface proteins. However, the three proteases do not all act on the same surface proteins; some proteins are more specifically cleaved by HtrA, while others are more specifically degraded by SepM (substrate-binding proteins of ABC transporters, for instance). Consequently, the three sheddases are not functionally redundant and instead act synergistically. In particular, SepM plays a significant role in the recycling of proteins involved in cell wall metabolism, and its inactivation has repercussions for cell morphology. In this respect, SepM could be considered a member of the bacterial cell morphogenesis orchestra (45).

In addition to this general recycling function—which was unknown until now for PrtS—PrtS and HtrA fulfill additional, specific roles. PrtS is responsible for extracellular protein degradation for nutritional purposes (4), while HtrA is involved in stress resistance, especially thermal stress (46). The role of SepM is not clear, at least in *S. thermophilus*. In pathogenic streptococci such as *Streptococcus agalactiae* or *Streptococcus gordonii*, SepM plays a role in virulence expression (47, 48), but this kind of role is not relevant in the GRAS species *S. thermophilus*. Additionally, SepM is reported to be involved in competence development in *S. mutans* and *S. gordonii* (22, 48). Here, however, analyses of streptide production and natural competence for transformation revealed no crucial role for SepM in the quorum sensing mechanisms of *S. thermophilus*. This finding should not be surprising considering the sequence of the pheromone that controls streptide production (MKKQILLTLLLVFEGIIIVVG); it does not contain Ala residues and thus should not be expected to be cleaved efficiently by SepM. Instead, the relationship between SepM and ComS, the competence-controlling pheromone, is more surprising. Based on the known specificity of SepM, the mature ComS₁₄₋₂₄ sequence (IALPYFAGCL) would be expected to be N-terminally cleaved between Ala and Ile. Although a previous study did detect a truncated form (ComS₁₆₋₂₄, ILPYFAGCL) in the supernatant of a ComS-overproducing strain, its abundance was lower than that of the mature form (40), suggesting that SepM acts only weakly on mature ComS₁₄₋₂₄. The situation is therefore somewhat different from that in *S. mutans*, where the precursor form of the CSP pheromone is activated through a proteolytic event—initiated by the ABC transporter dedicated to CSP export—and the resulting mature form of CSP (CSP-21) is produced in the supernatant (49). CSP-21 is then processed extracellularly into the CSP-18 form by the removal of three C-terminal amino acids; CSP-18 is much more active than CSP-21 in inducing competence, biofilm formation, and antimicrobial activity (50). Indeed, the CSP-18 form, which is released by SepM, corresponds to the biologically active form of the pheromone (22). Very surprisingly, our inactivation of SepM resulted in only a slight and statistically insignificant reduction of the competence efficiency of *S. thermophilus*. However, this result is consistent with work by Fontaine and colleagues, who compared the efficiency of competence induction of ComS₁₆₋₂₄ to that of other forms (ComS₁₇₋₂₄ and ComS₁₅₋₂₄) and found no significant differences (19). Therefore, it seems very unlikely that SepM from *S. thermophilus*

extracellularly cleaves ComS₁₄₋₂₄ into the more active ComS₁₆₋₂₄, and it can be concluded that SepM does not play a significant role in competence development.

The deletion of PrtS, HtrA, and SepM almost completely abolished cell surface proteolysis in *S. thermophilus*. This finding might have biotechnological applications and could be useful for the production of heterologous proteins. Among LABs, *L. lactis* is currently the main vector used for protein production (51), but *S. thermophilus* is considered a promising substitute (15). The ability to generate a strain deprived of shed-dase may increase the attractiveness of using *S. thermophilus* as a bacterial vector for protein production.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The bacterial strains used in this study are listed in Table 1. They were stored at -80°C in M17 broth supplemented with lactose (1%, wt/vol) and glycerol (20%, wt/vol) (52). Cultures were grown at 42°C in M17 supplemented with lactose (1%, wt/vol), in CDM (53), or in reconstituted skim milk (10%, wt/vol). When required, erythromycin (5 $\mu\text{g/ml}$), kanamycin (1 mg/ml), or spectinomycin (150 $\mu\text{g/ml}$) was added to M17 medium. When CDM cultures were used for exopeptidome determination (see below), concentrations of the aromatic amino acids Tyr, Phe, and Trp were reduced 10-fold; this reduction made it possible to increase the volume of the sample injected in the HPLC system without overloading the separation column. For streptide detection by HPLC, the CDM was prepared without aromatic amino acids and riboflavin, as riboflavin coeluted with streptide under our experimental conditions. For all other purposes, the composition of the CDM was not modified.

DNA manipulations and construction of mutant strains. Standard methods were used for DNA manipulations and sequencing. PCR was performed using the oligonucleotides listed in Table S4 in the supplemental material (Eurogentec, Seraing, Belgium). The single mutant strains LMD9 Δ htrA, LMD9 Δ prtS, and LMD9 Δ sepM were constructed using the overlapping PCR method (18). Briefly, the upstream and downstream regions of the genes to be inactivated were amplified by PCR using the appropriate oligonucleotides (Table S4), with chromosomal DNA of *S. thermophilus* LMD9 as a template. The erythromycin (*erm*), kanamycin (*aphA3*), and spectinomycin (*spec*) cassettes were amplified using the appropriate oligonucleotides and with plasmids pG+host9 (54), pKa (55), and pAT28 (56) as the templates, respectively. Upstream and downstream fragments of *htrA*, *prtS*, and *sepM* were fused by PCR to the *erm*, *aphA3*, and *spec* cassettes, respectively. The resulting fragments (2.41, 2.37, and 2.05 kb, respectively) were used to transform *S. thermophilus* LMD9 by natural transformation (40). Transformants were selected via plating on M17 medium that contained the appropriate antibiotic and checked by PCR. The final selected mutants were checked by sequencing. The double mutant Δ htrA Δ prtS was obtained by natural transformation of LMD9 Δ prtS with chromosomal DNA of LMD9 Δ htrA. The double mutants Δ htrA Δ sepM and Δ prtS Δ sepM and the triple mutant Δ htrA Δ prtS Δ sepM were obtained by natural transformation of LMD9 Δ htrA, LMD9 Δ prtS, and LMD9 Δ htrA Δ prtS, respectively, with the 2.05-kb PCR fragment corresponding to the upstream and downstream regions of *sepM* fused to the *spec* cassette. Transformants were selected via plating on M17 medium that contained the appropriate antibiotic and mutants were validated by PCR. The CNRZ1066 Δ htrA Δ sepM mutant strain was constructed using a similar strategy, except that chromosomal DNA of *S. thermophilus* CNRZ1066 was used as the template, and the natural competence of the strain was stimulated by adding 10 μM ComS₁₇₋₂₄ peptide (LPYFAGCL) 10 min prior to the addition of the PCR fragment (40).

Extracellular peptidome identification. Peptides accumulating in the culture medium of *S. thermophilus* LMD-9 and the various mutant strains were identified essentially as described previously (17). Briefly, strains were grown to an optical density at 600 nm (OD_{600}) of 1.0 in CDM that contained a 10-fold reduced concentration of aromatic amino acids. Cells were discarded by centrifugation ($5,000 \times g$ for 10 min at room temperature), and the supernatant was filtered through a 0.22- μm pore-size filter (low-binding polyvinylidene difluoride [PVDF] membrane; Millipore, Guyancourt, France). Peptides were concentrated by solid-phase extraction on StrataX cartridges (Phenomenex, Le Pecq, France) and the elution step was performed with 50% acetonitrile. Eluted peptides were dried in a SpeedVac (Savant; Thermo Scientific), resuspended in 0.1% trifluoroacetic acid (TFA), and ultrafiltered at 3 kDa (Centriplus YM-3; Millipore). An equivalent of 20 ml of supernatant was loaded on a reverse-phase C₁₈ column (Kinetex, 150 by 4.6 mm, 2.6 μm , 100 Å; Phenomenex). Peptides were separated at 40°C with a linear gradient of acetonitrile (1.6% per min) in ammonium formate (20 mM [pH 6.2]) at 0.7 ml/min. Sixteen fractions of 0.7 ml each were collected and dried.

Each of the 16 peptide-containing fractions was resuspended in 0.1% TFA and 2% acetonitrile and analyzed by LC-tandem MS (LC-MS/MS) at the PAPPSo platform (INRAE, Jouy-en-Josas). The sample used for peptidomic analysis was equivalent to 3.33 ml of culture medium. Peptide separation was performed on a C₁₈ column (Pepmap, 150 by 0.075 mm, 3 μm , 100 Å; Dionex, Thermo Scientific) at 300 nl/min with a linear gradient (0.8% per min) of acetonitrile in 0.1% formic acid over the course of 45 min. Eluted peptides were analyzed online using an LTQ-Orbitrap Discovery mass spectrometer (Thermo Fischer) as previously described (17). Peptide and protein identification were performed with X!Tandem version 2017.2.1.4 (Alanine) and X!Tandem pipeline version 3.4.3 (Elastine Durcie) (57); these analyses used the GenBank protein sequence of *S. thermophilus* LMD-9 (CP000419.1) (58) that is associated with a classical proteomic contaminant database. The X!Tandem search parameters included no enzymatic cleavage specificity, variable oxidation of methionine (+15.99491), acetylation of the N-terminal residue

(+42.01057), cyclization of the N-terminal Glu (−8.01056) or Gln (−17.02655), and mass tolerance of 10 ppm for the parent ion M+H mass and of 0.4 Da for the ion fragments. Peptide matches with E value of ≤ 0.01 were conserved (17, 59).

LC-MS/MS detection of streptide and SHP₁₃₅₈. The presence of streptide (AK*GDGW*KVM, with the asterisk indicating cyclization) and of the pheromone SHP₁₃₅₈ (EGLIVVVG) was assessed in culture media by mass spectrometry. *S. thermophilus* LMD9 and the protease mutant strains were grown in CDM to an OD₆₀₀ value of 1.0. Peptides present in the culture supernatant were separated on a Pepmap C₁₈ column at 300 nl/min with a linear gradient (1% per min) of acetonitrile in formic acid (0.1%) for 90 min. Eluted peptides were analyzed online on an LTQ-Orbitrap Discovery mass spectrometer (Thermo Fischer) using the same experimental separation conditions as for the peptidomic analyses. Data were analyzed using Qual Browser Thermo Xcalibur version 2.2 (Thermo Scientific). The ion current signals of the doubly charged streptide ($[M+H]^{2+} = 495.2474$) and the monocharged pheromone SHP₁₃₅₈ ($[M+H]^+ = 1018.5608$) were extracted, and peptide identification was validated by analyzing the fragmentation spectra. The total area of each peak was calculated, giving an estimation of the relative abundance of the peptides.

HPLC detection of streptide. Cells were grown overnight in CDM deprived of aromatic amino acids and riboflavin. After cells were discarded by centrifugation, TFA (0.1%, vol/vol) and acetonitrile (2%, vol/vol) were added to the supernatant. Streptide was concentrated by solid-phase extraction using a StrataX polymeric reversed phase cartridge (Phenomenex) according to the manufacturer's instructions. Thirty milliliters of sample were loaded on the column, and elution was performed with a 20% acetonitrile solution. The eluted solution was dried in a Speed Vac, resuspended in 0.1% TFA and 2% acetonitrile with a 300-fold concentration factor, and ultrafiltered through a 3-kDa ultrafiltration membrane (Centriplus YM-3; Millipore). Fifty microliters of the resuspended solution were loaded onto a Kinetex C₁₈ HPLC column (150 by 4.6 mm, 2.6 μ m 100 Å; Phenomenex). Streptide was eluted using a linear gradient of 5% to 30% acetonitrile in 0.1% TFA over 25 min and detected at 214 nm.

Scanning electron microscopy. Bacteria grown in CDM to the exponential (OD₆₀₀ 0.8) or stationary phase were harvested by centrifugation ($5,000 \times g$ for 10 min at room temperature) and washed in 50 mM KH₂PO₄-K₂HPO₄ (pH 6.7). Washed cells were fixed, desiccated, dried, and metallized as described previously (60). Sample pictures were acquired and analyzed at the MIMA2 microscopy and imaging platform (INRAE; <http://doi.org/10.15454/1.5572348210007727E12>) using a Hitachi SU5000 field electron gun (FEG) low vacuum scanning electron microscope with secondary electron (SE) detector operating at 2 kV.

Miscellaneous. Peptidome contents (small independent samples) and peak areas of streptide and SHP₁₃₅₈ (small paired samples) from wild-type and mutant strains were compared statistically using the nonparametric Kruskal-Wallis rank sum test and the nonparametric paired Wilcoxon test, respectively, using the coin package of R (version 3.2.1). Protein alignments were performed using version 2.9.0 of BLASTP (61). The 2D structure of SepM was predicted using the PSIPRED protein structure prediction tool (<http://bioinf.cs.ucl.ac.uk/psipred>) (29). The synteny of *sepM* flanking genes was obtained from the SyntTax Web server (<https://archaea.i2bc.paris-saclay.fr/SyntTax>).

Data availability. The mass spectrometry peptidomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (62) partner repository with the identifiers PXD027660 (wild-type strain), PXD027643 (*prtS* mutant), PXD027644 (*htrA* mutant), PXD027666 (*sepM* mutant), PXD027665 (*prtS htrA* double mutant), and PXD027671 (*prtS htrA sepM* triple mutant).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 1.3 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.03 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.03 MB.

SUPPLEMENTAL FILE 4, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

We are grateful to the INRAE PAPPISO proteomics platform (<http://pappiso.inra.fr/en/>), which is supported by INRAE (<http://www.inrae.fr>), the Ile-de-France regional council (<https://www.iledefrance.fr/education-recherche>), IBISA (<https://www.ibisa.net>), and CNRS (<http://www.cnrs.fr>) for providing mass spectrometry facilities and the INRAE MIGALE bioinformatics platform (<http://migale.inrae.fr>) for providing computational resources. We warmly thank Marie-Pierre Chapot-Chartier and Saulius Kulakauskas for fruitful discussions and Françoise Rul and Christine Delorme for critical reading of the manuscript.

We declare no conflict of interest.

V.M., R.G., and V.J. contributed to the conception and the design of the study; M.B. and L.O.C. performed peptidomics analysis; M.B., C. Metton, and C. Mézange performed mutant constructions; T.M. performed MEB analyses; M.B., C. Metton, T.M., R.G., and V.J. contributed to the analysis and the interpretation of the data; V.M., R.G., and V.J.

contributed to the writing of the manuscript. The authors have all read and approved the final version of the manuscript.

REFERENCES

- Delorme C. 2008. Safety assessment of dairy microorganisms: *Streptococcus thermophilus*. *Int J Food Microbiol* 126:274–277. <https://doi.org/10.1016/j.ijfoodmicro.2007.08.014>.
- Iyer R, Tomar SK, Maheswari TU, Singh R. 2010. *Streptococcus thermophilus* strains: multifunctional lactic acid bacteria. *Int Dairy J* 20:133–141. <https://doi.org/10.1016/j.idairyj.2009.10.005>.
- Kunji ERS, Mierau I, Hagting A, Poolman B, Konings WN. 1996. The proteolytic systems of lactic acid bacteria. *Antonie Van Leeuwenhoek* 70: 187–221. <https://doi.org/10.1007/BF00395933>.
- Fernandez-Espla MD, Garault P, Monnet V, Rul F. 2000. *Streptococcus thermophilus* cell wall-anchored proteinase: release, purification, and biochemical and genetic characterization. *Appl Environ Microbiol* 66: 4772–4778. <https://doi.org/10.1128/AEM.66.11.4772-4778.2000>.
- de Vos WM, Vos P, de Haard H, Boerrigter I. 1989. Cloning and expression of the *Lactococcus lactis* subsp. *cremoris* SK11 gene encoding an extracellular serine proteinase. *Gene* 85:169–176. [https://doi.org/10.1016/0378-1119\(89\)90477-0](https://doi.org/10.1016/0378-1119(89)90477-0).
- Wang BY, Kuramitsu HK. 2005. Interactions between oral bacteria: inhibition of *Streptococcus mutans* bacteriocin production by *Streptococcus gordonii*. *Appl Environ Microbiol* 71:354–362. <https://doi.org/10.1128/AEM.71.1.354-362.2005>.
- Poquet I, Saint V, Seznec E, Simoes N, Bolotin A, Gruss A. 2000. HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing. *Mol Microbiol* 35:1042–1051. <https://doi.org/10.1046/j.1365-2958.2000.01757.x>.
- Clausen T, Kaiser M, Huber R, Ehrmann M. 2011. HTRA proteases: regulated proteolysis in protein quality control. *Nat Rev Mol Cell Biol* 12: 152–162. <https://doi.org/10.1038/nrm3065>.
- Liu Y, Zeng Y, Huang Y, Gu L, Wang S, Li C, Morrison DA, Deng H, Zhang JR. 2019. HtrA-mediated selective degradation of DNA uptake apparatus accelerates termination of pneumococcal transformation. *Mol Microbiol* 112:1308–1325. <https://doi.org/10.1111/mmi.14364>.
- Cassone M, Gagne AL, Spruce LA, Seeholzer SH, Seibert ME. 2012. The HtrA protease from *Streptococcus pneumoniae* digests both denatured proteins and the competence-stimulating peptide. *J Biol Chem* 287: 38449–38459. <https://doi.org/10.1074/jbc.M112.391482>.
- Kochan TJ, Dawid S. 2013. The HtrA protease of *Streptococcus pneumoniae* controls density-dependent stimulation of the bacteriocin blp locus via disruption of pheromone secretion. *J Bacteriol* 195:1561–1572. <https://doi.org/10.1128/JB.01964-12>.
- Markakiou S, Gaspar P, Johansen E, Zeidan AA, Neves AR. 2020. Harnessing the metabolic potential of *Streptococcus thermophilus* for new biotechnological applications. *Curr Opin Biotechnol* 61:142–152. <https://doi.org/10.1016/j.copbio.2019.12.019>.
- van de Guchte M, Kok J, Venema G. 1992. Gene expression in *Lactococcus lactis*. *FEMS Microbiol Rev* 8:73–92. <https://doi.org/10.1111/j.1574-6968.1992.tb04958.x>.
- Nouaille S, Ribeiro LA, Miyoshi A, Pontes D, Le Loir Y, Oliveira SC, Langella P, Azevedo V. 2003. Heterologous protein production and delivery systems for *Lactococcus lactis*. *Genet Mol Res* 2:102–111.
- Lecomte X, Gagnaire V, Lortal S, Dary A, Genay M. 2016. *Streptococcus thermophilus*, an emerging and promising tool for heterologous expression: advantages and future trends. *Food Microbiol* 53:2–9. <https://doi.org/10.1016/j.fm.2015.05.003>.
- Morello E, Bermudez-Humaran LG, Lull D, Solé V, Miraglio N, Langella P, Poquet I. 2008. *Lactococcus lactis*, an efficient cell factory for recombinant protein production and secretion. *J Mol Microbiol Biotechnol* 14:48–58. <https://doi.org/10.1159/000106082>.
- Guillot A, Boulay M, Chambellon E, Gitton C, Monnet V, Juillard V. 2016. Mass spectrometry analysis of the extracellular peptidome of *Lactococcus lactis*: lines of evidence for the coexistence of extracellular protein hydrolysis and intracellular peptide excretion. *J Proteome Res* 15:3214–3224. <https://doi.org/10.1021/acs.jproteome.6b00424>.
- Gardan R, Besset C, Guillot A, Gitton C, Monnet V. 2009. The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J Bacteriol* 191:4647–4655. <https://doi.org/10.1128/JB.00257-09>.
- Fontaine L, Boutry C, de Frahan MH, Delplace B, Fremaux C, Horvath P, Boyaval P, Hols P. 2010. A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J Bacteriol* 192:1444–1454. <https://doi.org/10.1128/JB.01251-09>.
- Mulder J, Wels M, Kuipers OP, Kleerebezem M, Bron PA. 2017. Unleashing natural competence in *Lactococcus lactis* by induction of the competence regulator ComX. *Appl Environ Microbiol* 83:e01320-17. <https://doi.org/10.1128/AEM.01320-17>.
- David B, Radziejewski A, Toussaint F, Fontaine L, Henry de Frahan M, Patout C, van Dillen S, Boyaval P, Horvath P, Fremaux C, Hols P. 2017. Natural DNA transformation is functional in *Lactococcus lactis* subsp. *cremoris* KW2. *Appl Environ Microbiol* 83:e01074-17. <https://doi.org/10.1128/AEM.01074-17>.
- Hossain MS, Biswas I. 2012. An extracellular protease, SepM, generates functional competence-stimulating peptide in *Streptococcus mutans* UA159. *J Bacteriol* 194:5886–5896. <https://doi.org/10.1128/JB.01381-12>.
- Zhou M, Boekhorst J, Francke C, Siezen RJ. 2008. LocateP: genome-scale subcellular localization predictor for bacterial proteins. *BMC Bioinformatics* 9:173. <https://doi.org/10.1186/1471-2105-9-173>.
- Wang X, Zhang J, Li G-Z. 2015. Multi-location gram-positive and gram-negative bacterial protein subcellular localization using gene ontology and multi-label classifier ensemble. *BMC Bioinformatics* 16:S1. <https://doi.org/10.1186/1471-2105-16-S12-S1>.
- Rawlings ND, Barret AJ, Finn R. 2016. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 44:D343–D350. <https://doi.org/10.1093/nar/gkv1118>.
- Courtin P, Miranda G, Guillot A, Wessner F, Mézange C, Domakova E, Kulakauskas S, Chapot-Chartier M-P. 2006. Peptidoglycan structure analysis of *Lactococcus lactis* reveals the presence of an α , β -carboxypeptidase involved in peptidoglycan maturation. *J Bacteriol* 188:5293–5298. <https://doi.org/10.1128/JB.00285-06>.
- Shahbal S, Hemme D, Renault P. 1993. Characterization of a cell-envelope-associated proteinase activity from *Streptococcus thermophilus* H-strains. *Appl Environ Microbiol* 59:177–182. <https://doi.org/10.1128/aem.59.1.177-182.1993>.
- Buchan DWA, Jones DT. 2019. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res* 47:W402–W407. <https://doi.org/10.1093/nar/gkz297>.
- Sillitoe I, Bordin N, Dawson N, Waman VP, Ashford P, Scholes HM, Pang CSM, Woodridge L, Rauer C, Sen N, Abbasian M, Le Cornu S, Lam SD, Berka K, Varekova IH, Svobodova R, Lees J, Orenge CA. 2020. CATH: increased structural coverage of functional space. *Nucleic Acids Res* 49: D266–D273. <https://doi.org/10.1093/nar/gkaa1079>.
- Botos I, Melnikov EE, Cherry S, Tropea JE, Khalatova AG, Rasulova F, Dauter Z, Maurizi MR, Rotanova TV, Wlodawer A, Gustchina A. 2004. The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys Dyad in the active site. *J Biol Chem* 279:8140–8148. <https://doi.org/10.1074/jbc.M312243200>.
- Caspi R, Billington R, Fulcher CA, Keseler WK, Kothari A, Krummenacker M, Latendresse M, Midford PE, Ong Q, Ong WK, Paley S, Subhraveti P, Karp PD. 2018. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* 46:D633–D639. <https://doi.org/10.1093/nar/gkx935>.
- Lingeswaran A, Metton C, Henry C, Monnet V, Juillard V, Gardan R. 2020. Export of Rgg quorum sensing peptides is mediated by the PptAB ABC transporter in *Streptococcus thermophilus* strain LMD-9. *Genes* 11:1096. <https://doi.org/10.3390/genes11091096>.
- Proust L, Haudebourg E, Sourabié A, Pedersen M, Besançon I, Monnet V, Juillard V. 2020. Multi-omics approach reveals how yeast extract peptides shape *Streptococcus thermophilus* metabolism. *Appl Environ Microbiol* 86: e01446-20. <https://doi.org/10.1128/AEM.01446-20>.
- Eng C, Thibessard A, Danielsen M, Rasmussen TB, Mari J-F, Leblond P. 2011. *In silico* prediction of horizontal gene transfer in *Streptococcus thermophilus*. *Arch Microbiol* 193:287–297. <https://doi.org/10.1007/s00203-010-0671-8>.

35. Ricomini Filho AP, Khan R, Åmdal HA, Petersen FC. 2019. Conserved pheromone production, response and degradation by *Streptococcus mutans*. *Front Microbiol* 10:2140. <https://doi.org/10.3389/fmicb.2019.02140>.
36. Ibrahim M, Guillot A, Wessner F, Algaron F, Besset C, Courtin P, Gardan R, Monnet V. 2007. Control of the transcription of a short gene encoding a peptide in *Streptococcus thermophilus*: a new quorum sensing system? *J Bacteriol* 189:8844–8854. <https://doi.org/10.1128/JB.01057-07>.
37. Fleuchot B, Gitton C, Guillot A, Vidic J, Nicolas P, Besset C, Fontaine L, Hols P, Leblond-Bourget N, Monnet V, Gardan R. 2011. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol Microbiol* 80:1102–1119. <https://doi.org/10.1111/j.1365-2958.2011.07633.x>.
38. Fontaine L, Boutry C, Guédon E, Guillot A, Ibrahim M, Grossiord B, Hols P. 2007. Quorum-sensing regulation of the production of Blp bacteriocins in *Streptococcus thermophilus*. *J Bacteriol* 189:7195–7205. <https://doi.org/10.1128/JB.00966-07>.
39. Schramma KR, Bushin LB, Seyedsayamdost MR. 2015. Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. *Nat Chem* 7:431–437. <https://doi.org/10.1038/nchem.2237>.
40. Gardan R, Besset C, Gitton C, Guillot A, Fontaine L, Hols P, Monnet V. 2013. Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. *J Bacteriol* 195:1845–1855. <https://doi.org/10.1128/JB.02196-12>.
41. Dallas DC, Guerrero A, Parker EA, Robinson RC, Gan J, German JB, Barile D, Lebrilla CB. 2015. Current peptidomics: applications, purification, identification, quantification and functional analysis. *Proteomics* 15:1026–1038. <https://doi.org/10.1002/pmic.201400310>.
42. Gasson M. 1983. Plasmid complements of *Streptococcus lactis* NCDO712 and other lactic streptococci after protoplast-induced curing. *J Bacteriol* 154:1–9. <https://doi.org/10.1128/jb.154.1.1-9.1983>.
43. Hedstrom L. 2002. Serine protease mechanism and specificity. *Chem Rev* 102:4501–4523. <https://doi.org/10.1021/cr000033x>.
44. Biswas S, Cao L, Kim A, Biswas I. 2016. SepM, a streptococcal protease involved in quorum sensing, displays strict substrate specificity. *J Bacteriol* 198:436–447. <https://doi.org/10.1128/JB.00708-15>.
45. Carballido-Lopez R. 2006. Orchestrating bacterial cell morphogenesis. *Mol Microbiol* 60:815–819. <https://doi.org/10.1111/j.1365-2958.2006.05161.x>.
46. Clausen T, Southan C, Ehrmann M. 2002. The HtrA family of proteases: implications for protein composition and cell fate. *Mol Cell* 10:443–455. [https://doi.org/10.1016/S1097-2765\(02\)00658-5](https://doi.org/10.1016/S1097-2765(02)00658-5).
47. Johri AK, Margarit I, Broenstrup M, Brettoni C, Hua L, Gygi SP, Telford JL, Grandi G, Paoletti LC. 2007. Transcriptional and proteomic profiles of Group B *Streptococcus* type V reveal potential adherence proteins associated with high-level invasion. *Infect Immun* 75:1473–1483. <https://doi.org/10.1128/IAI.00638-06>.
48. Pavlova SI, Wilkening RV, Federle MJ, Lu Y, Schwartz J, Tao L. 2019. *Streptococcus* endopeptidases promote HPV infection in vitro. *MicrobiologyOpen* 8:e00628. <https://doi.org/10.1002/mbo3.628>.
49. Hale JD, Heng NC, Jack RW, Tagg JR. 2005. Identification of nlmTE, the locus encoding the ABC transport system required for export of nonantibiotic mutacins in *Streptococcus mutans*. *J Bacteriol* 187:5036–5039. <https://doi.org/10.1128/JB.187.14.5036-5039.2005>.
50. Petersen FC, Fimland G, Scheie AA. 2006. Purification and functional studies of a potent modified quorum-sensing peptide and a two-peptide bacteriocin in *Streptococcus mutans*. *Mol Microbiol* 61:1322–1334. <https://doi.org/10.1111/j.1365-2958.2006.05312.x>.
51. Song AAL, In LLA, Lim SHE, Rahim RA. 2017. A review on *Lactococcus lactis*: from food to factory. *Microb Cell Fact* 16:55. <https://doi.org/10.1186/s12934-017-0669-x>.
52. Terzaghi BE, Sandine WE. 1975. Improved medium for lactic streptococci and their bacteriophages. *Appl Microbiol* 29:807–813. <https://doi.org/10.1128/am.29.6.807-813.1975>.
53. Letort C, Juillard V. 2001. Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus*. *J Appl Microbiol* 91:1023–1029. <https://doi.org/10.1046/j.1365-2672.2001.01469.x>.
54. Biswas I, Gruss A, Ehrlich SD, Maguin E. 1993. High-efficiency gene inactivation and replacement system for gram-positive bacteria. *J Bacteriol* 175:3628–3635. <https://doi.org/10.1128/jb.175.11.3628-3635.1993>.
55. Trieu-Cuot P, Courvalin P. 1983. Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'5'-aminoglycoside phosphotransferase type III. *Gene* 23:331–341. [https://doi.org/10.1016/0378-1119\(83\)90022-7](https://doi.org/10.1016/0378-1119(83)90022-7).
56. Trieu-Cuot P, Carlier C, Poyart-Salmeron C, Courvalin P. 1990. A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Res* 18:4296. <https://doi.org/10.1093/nar/18.14.4296>.
57. Langella O, Valot B, Balliau T, Blein-Nicolas M, Bonhomme L, Zivy M. 2017. XITandem Pipeline: a tool to manage sequence redundancy for protein inference and phosphosite identification. *J Proteome Res* 16:494–503. <https://doi.org/10.1021/acs.jproteome.6b00632>.
58. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee J-H, Diaz-Muñiz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozayavkin S, Weimer B, Mills D. 2006. Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* 103:15611–15616. <https://doi.org/10.1073/pnas.0607117103>.
59. Guerrero A, Dallas DC, Contreras S, Chee S, Parker EA, Sun X, Dimapasoc L, Barile D, German JB, Lebrilla C. 2014. Mechanistic peptidomics: factors that dictate specificity in the formation of endogenous peptides in human milk. *Mol Cell Proteomics* 13:3343–3351. <https://doi.org/10.1074/mcp.M113.036194>.
60. Oxaran V, Ledue-Clier F, Dieye Y, Herry J-M, Péchoux C, Meylheuc T, Briandet R, Juillard V, Piard J-C. 2012. Pilus biogenesis in *Lactococcus lactis*. Molecular characterization and role in aggregation and biofilm formation. *PLoS One* 7:e50989. <https://doi.org/10.1371/journal.pone.0050989>.
61. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schäffer AA, Yu Y-K. 2005. Protein databases searches using compositionally adjusted substitution matrices. *FEBS J* 272:5101–5109. <https://doi.org/10.1111/j.1742-4658.2005.04945.x>.
62. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AK, Ternent T, Brazma A, Vizcaino JA. 2019. The PRIDE database and related tools in2019: improving support for quantification data. *Nucleic Acids Res* 47:D442–D450. <https://doi.org/10.1093/nar/gky1106>.
63. Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goltsman E, Mazur M, Pusch GD, Fonstein M, Overbeek R, Kyprides N, Purnelle B, Prozzi D, Ngui K, Masuy D, Hancy F, Berteau S, Boutry M, Delcour J, Goffeau A, Hols P. 2004. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* 22:1554–1558. <https://doi.org/10.1038/nbt1034>.